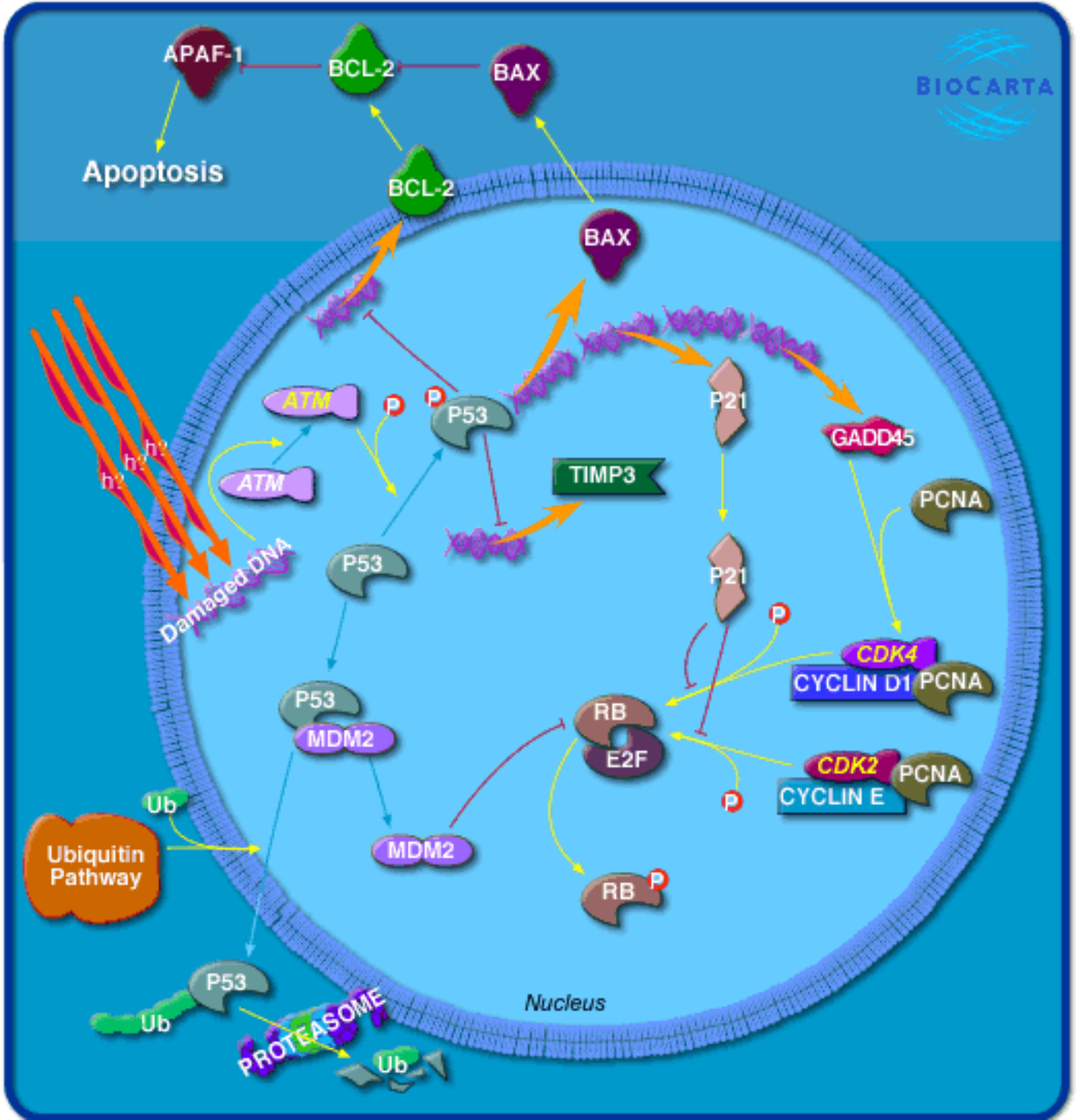
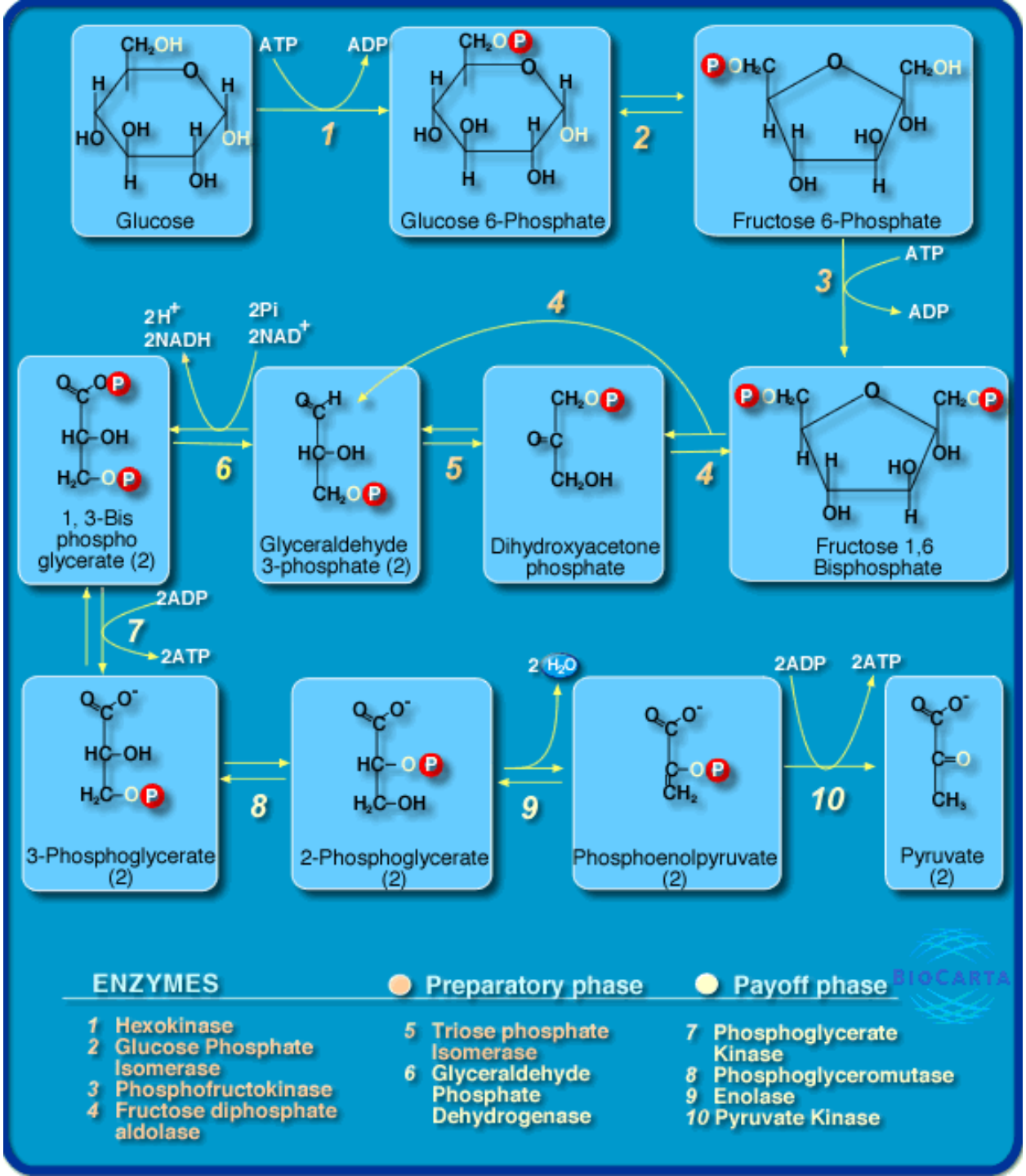
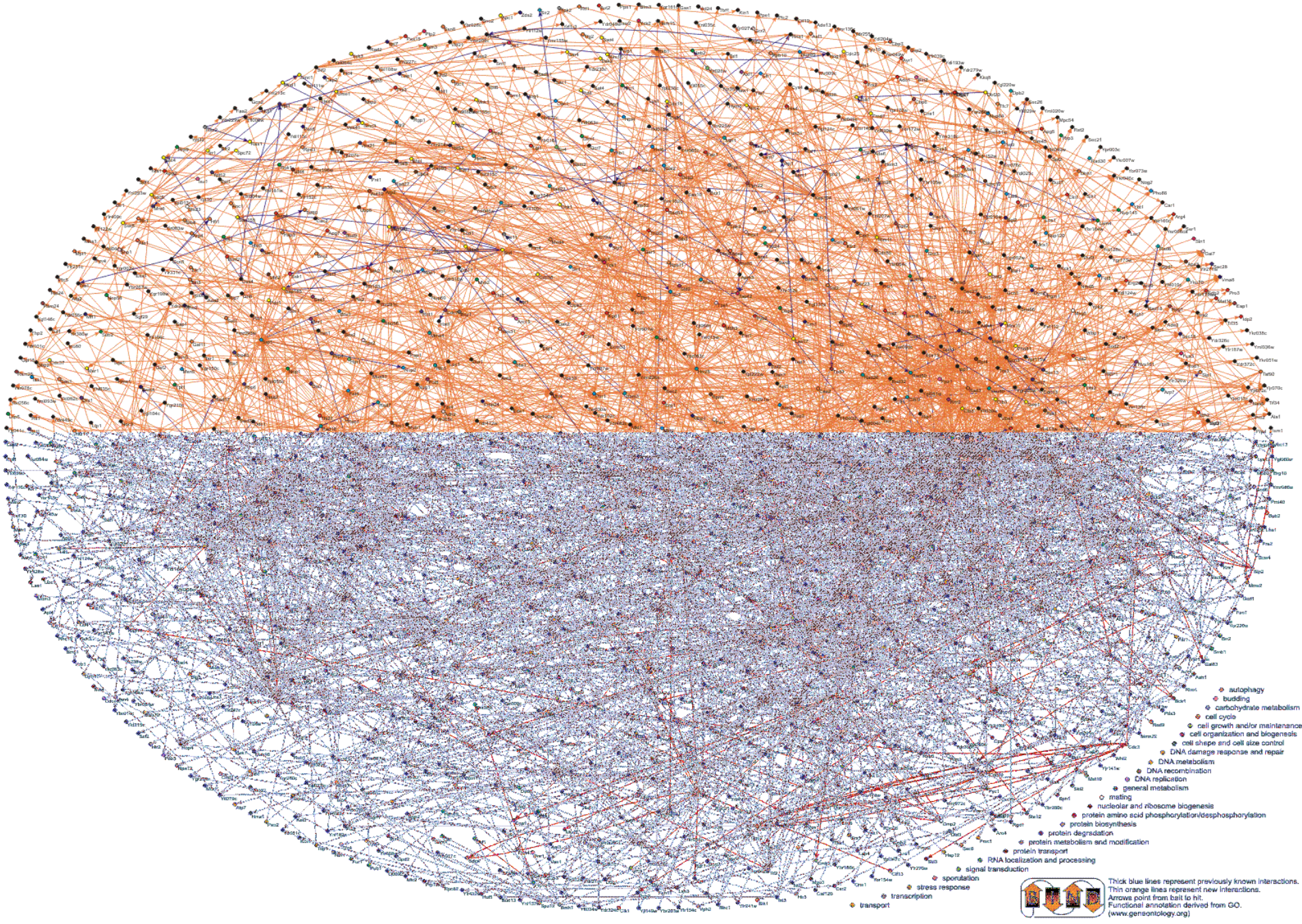
BioPAX
A Data Exchange Format for Biological Pathways

BioPAX Workgroup
www.biopax.org
caBIG July 2004


APAF-1
BCL-2
BAX
Apoptosis
BCL-2
BAX
BIOCARTA
ATM
P
P
P53
ATM
TIMP3
P21
GADD45
PCNA
h?
h?
h?
Damaged DNA
P53
P21
P
CDK4
CYCLIN D1
PCNA
P53
MDM2
RB
E2F
CDK2
PCNA
CYCLIN E
P
Ub
Ubiquitin
Pathway
MDM2
RB
P
P53
Ub
PROTEASOME
Ub
Nucleus


Glucose
ATP
ADP
1
Glucose 6-Phosphate
2
Fructose 6-Phosphate
ATP
3
ADP
2H+
2NADH
2Pi
2NAD+
4
1, 3-Bis phospho glycerate (2)
6
Glyceraldehyde 3-phosphate (2)
5
Dihydroxyacetone phosphate
4
Fructose 1,6 Bisphosphate
2ADP
7
2ATP
3-Phosphoglycerate (2)
8
2-Phosphoglycerate (2)
2 H2O
9
Phosphoenolpyruvate (2)
2ADP
2ATP
10
Pyruvate (2)
ENZYMES
Preparatory phase
Payoff phase
1 Hexokinase
2 Glucose Phosphate Isomerase
3 Phosphofructokinase
4 Fructose diphosphate aldolase
5 Triose phosphate Isomerase
6 Glyceraldehyde Phosphate Dehydrogenase
7 Phosphoglycerate Kinase
8 Phosphoglyceromutase
9 Enolase
10 Pyruvate Kinase
BIOCARTA

Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry



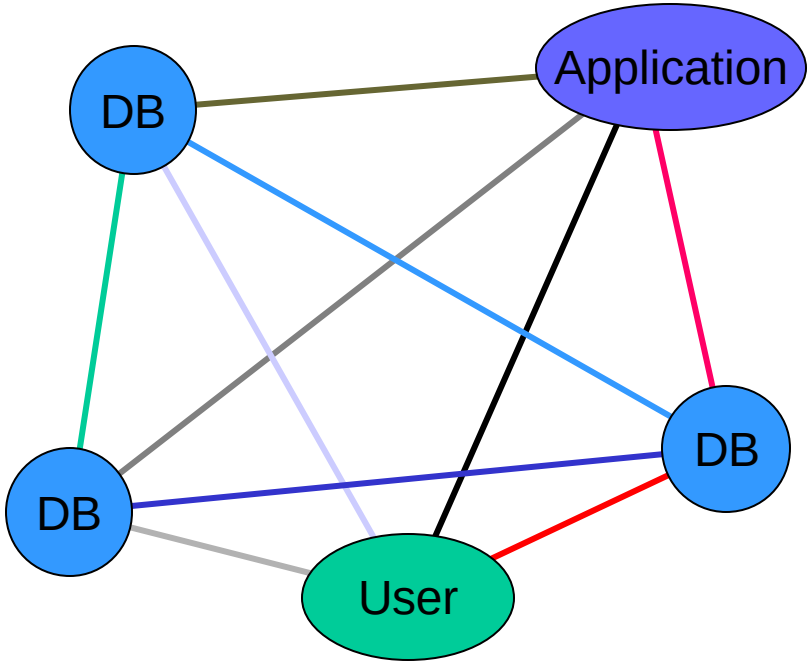
BioPAX Goals

- BioPAX = Biological Pathway Exchange
- Data exchange format for pathway data
- Include support for these pathway types:
  - Metabolic pathways
  - Signaling pathways
  - Protein-protein interactions
  - Genetic regulatory pathways
- Accommodate representations used in existing databases such as BioCyc, BIND, WIT, aMAZE, KEGG, etc.

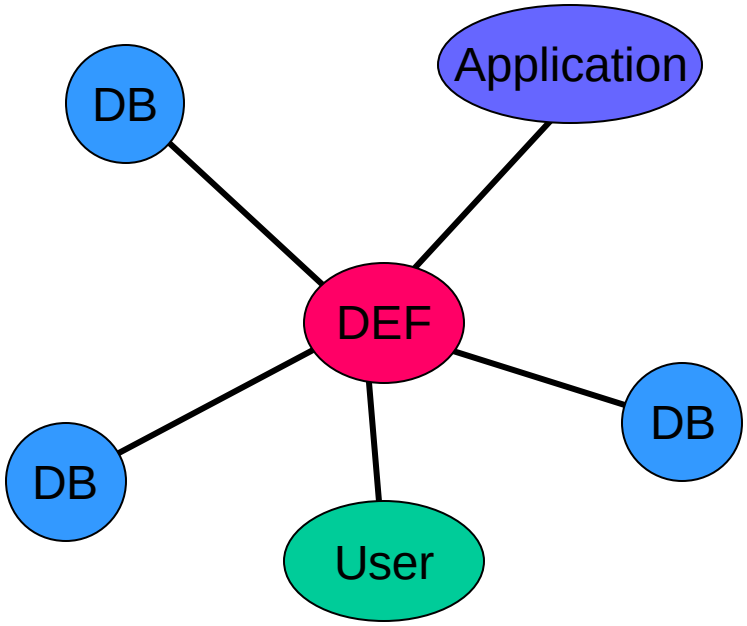
BioPAX Motivation

Common good

Before DEF



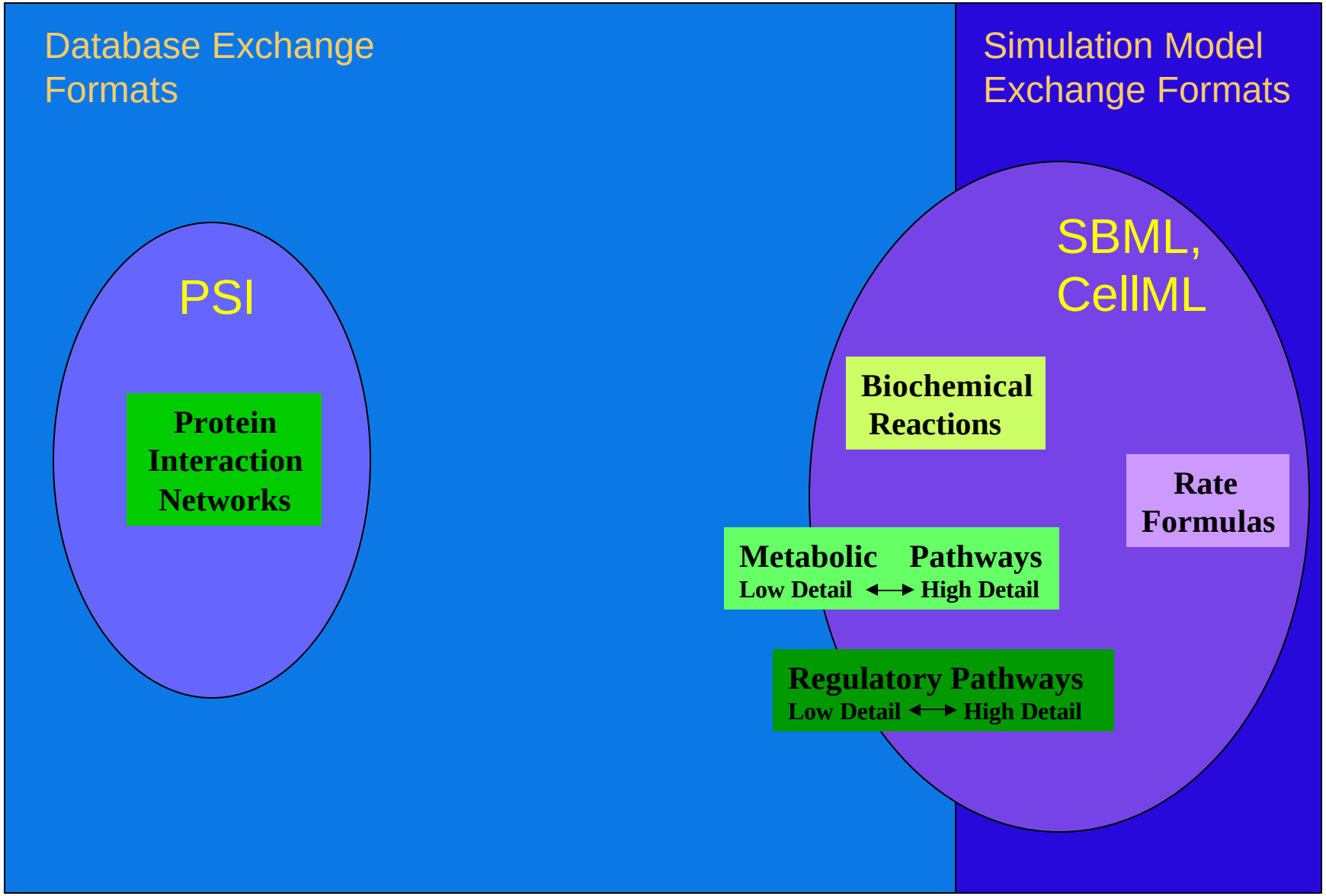
After DEF



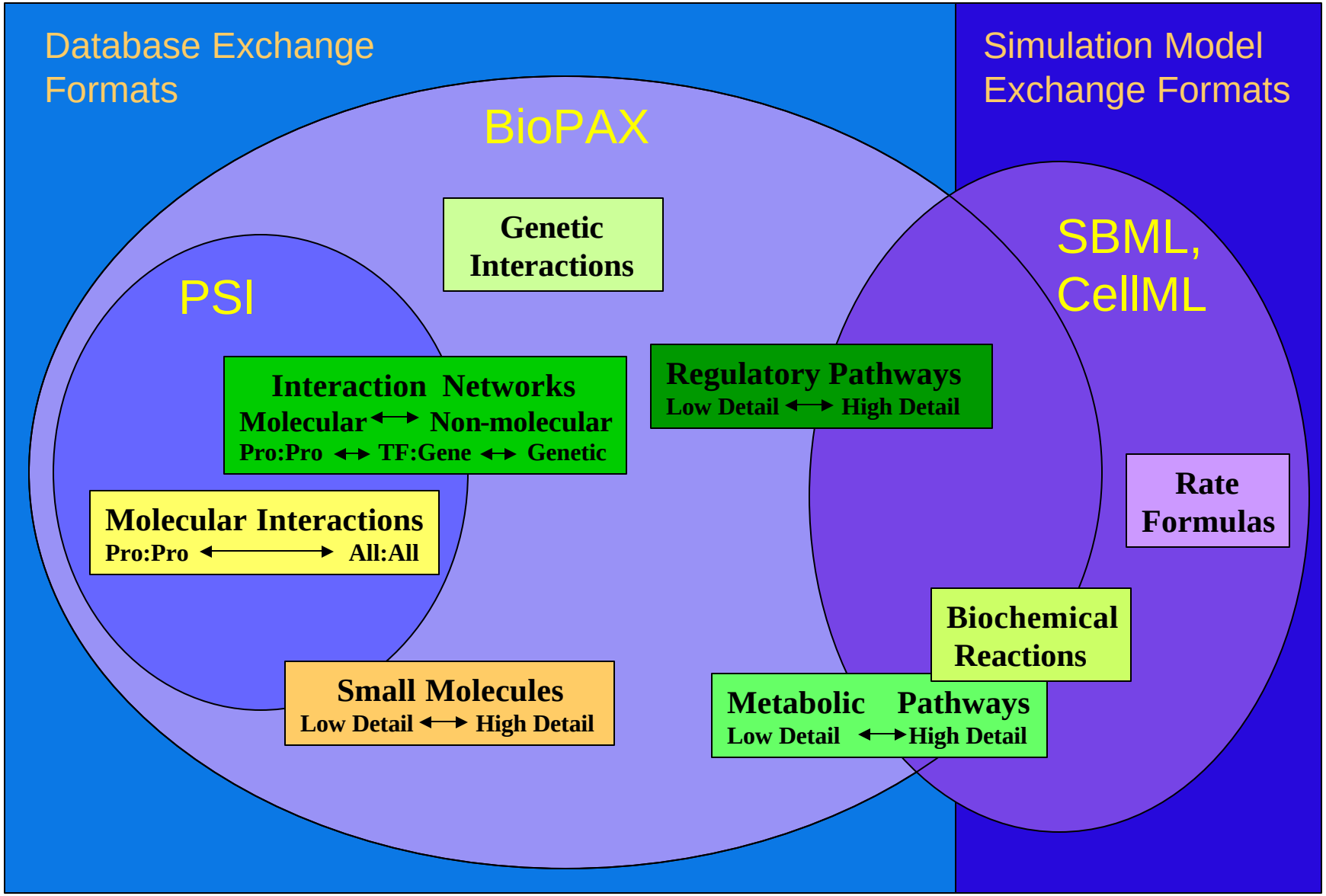
>100 DBs and tools

Promotes collaboration (big science), accessibility

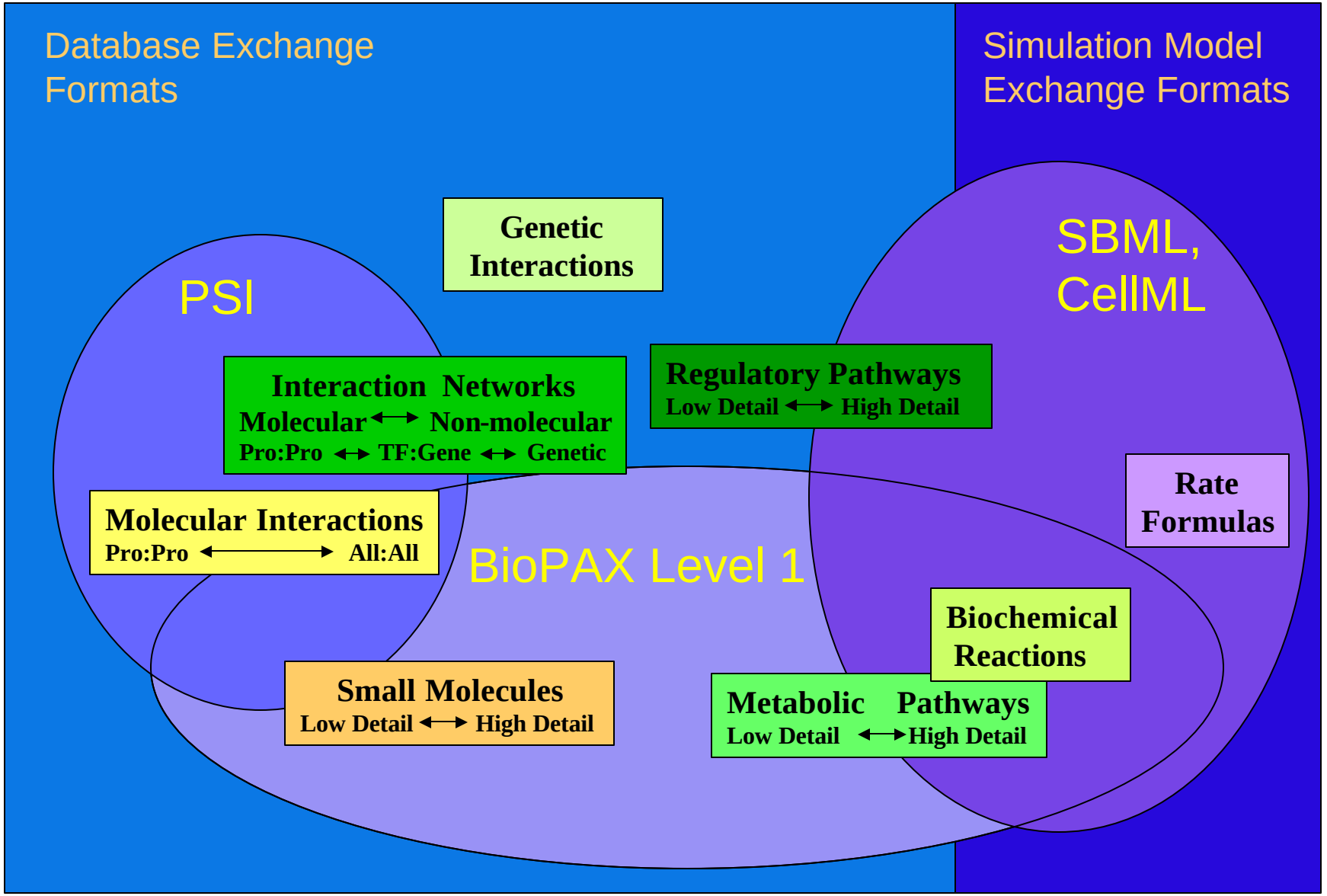
Exchange Formats in the Pathway Data Space



Exchange Formats in the Pathway Data Space



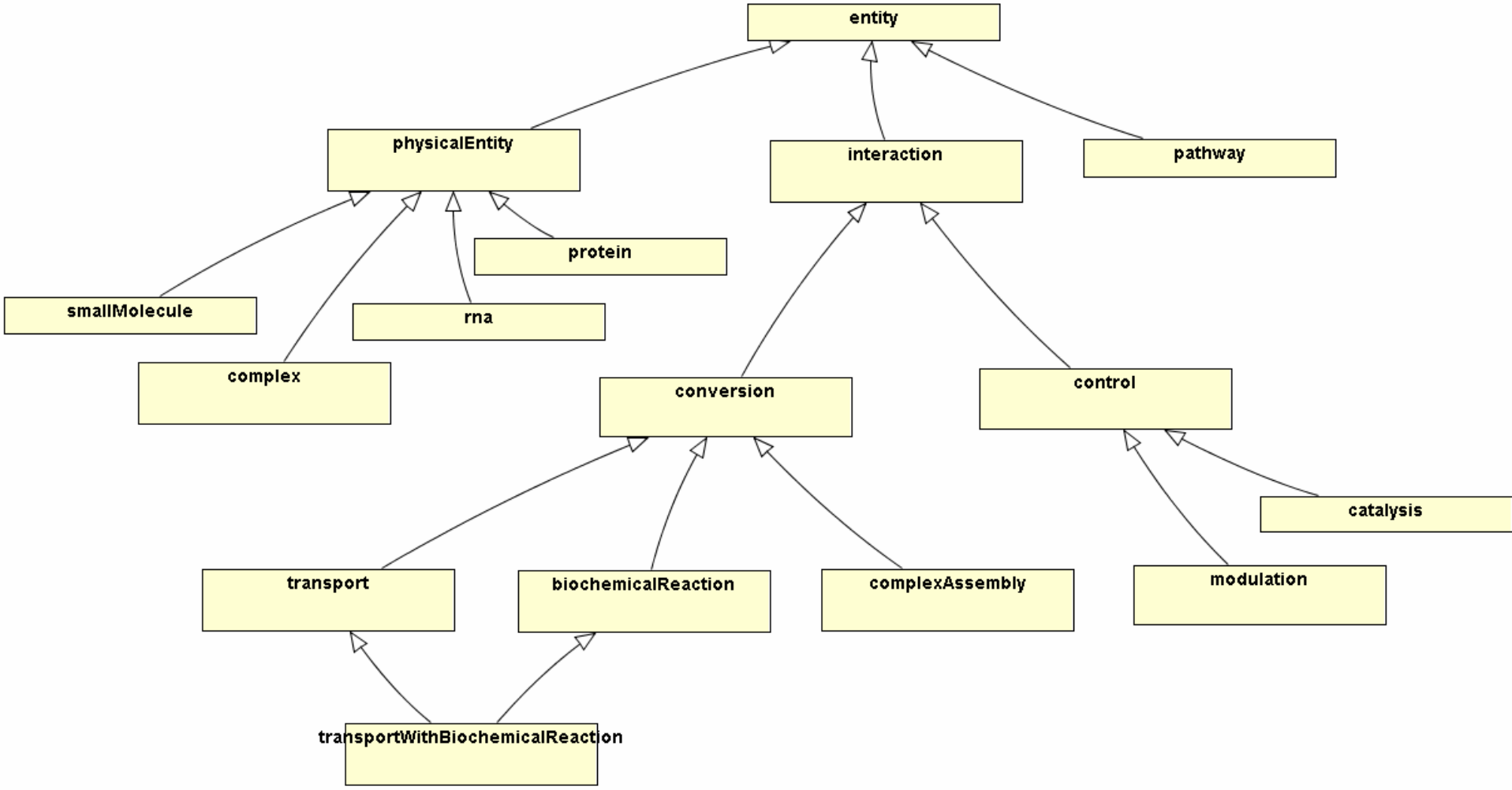
Exchange Formats in the Pathway Data Space



BioPAX

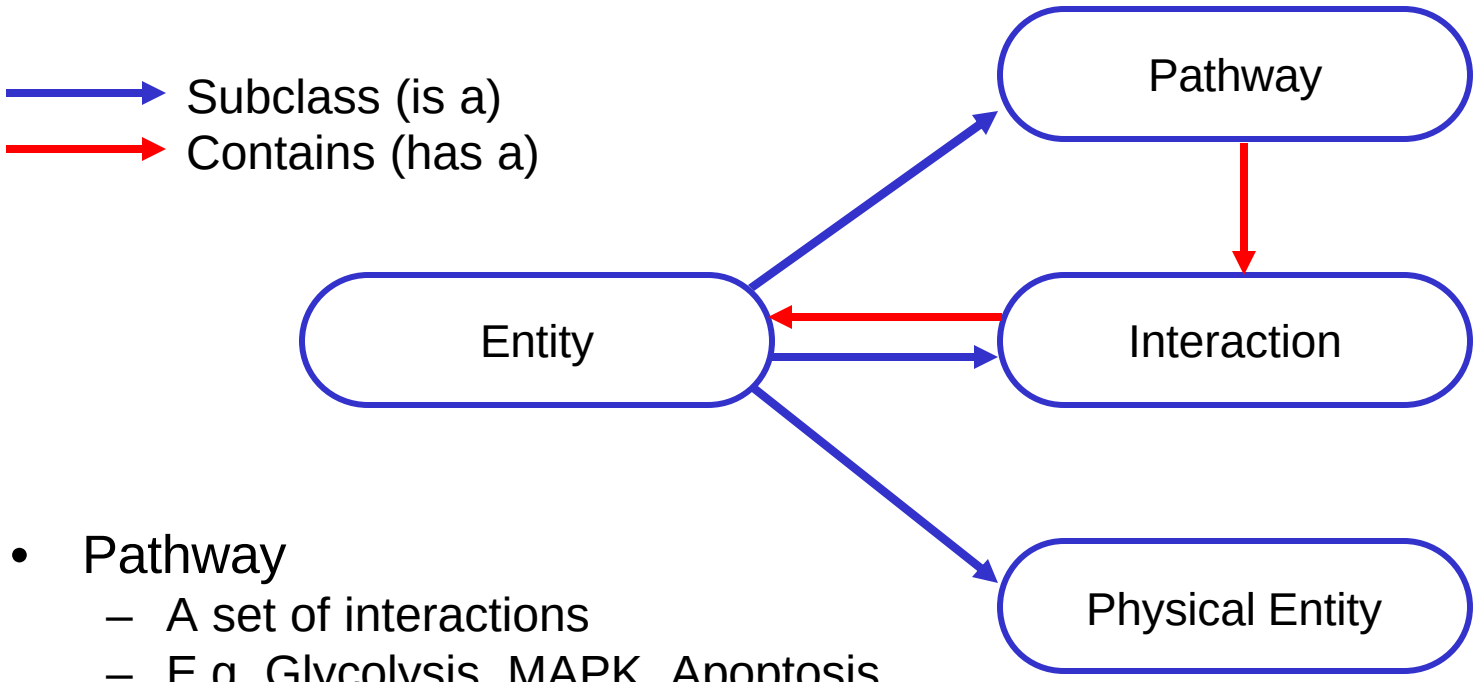
- Conceptual framework based upon existing DB schemas:
  - aMAZE, BIND, EcoCyc, WIT, KEGG, etc.
  - Allows wide range of detail, multiple levels of abstraction
- BioPAX ontology and format in OWL (XML)
- Ontology built using GKB Editor and Protégé
- Level 1 represents metabolic pathway data

BioPAX Ontology: Overview



Level 1 v1.0 (July 6th, 2004)

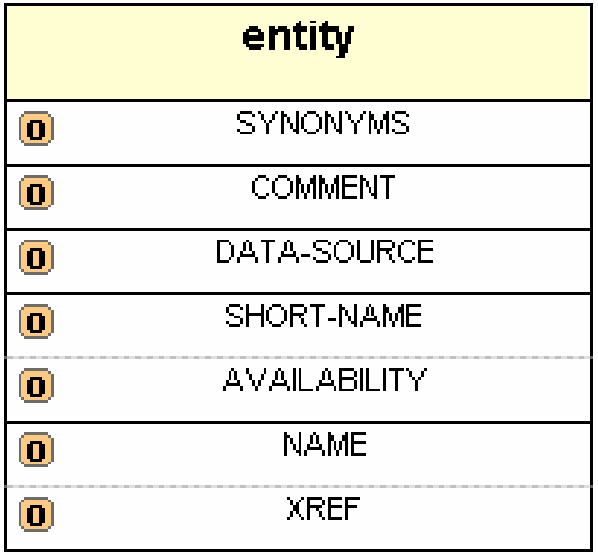
BioPAX Ontology: Top Level



- Pathway
  - A set of interactions
  - E.g. Glycolysis, MAPK, Apoptosis
- Interaction
  - A set of entities and some relationship between them
  - E.g. Reaction, Molecular Association, Catalysis
- Physical Entity
  - A building block of simple interactions
  - E.g. Small molecule, Protein, DNA, RNA

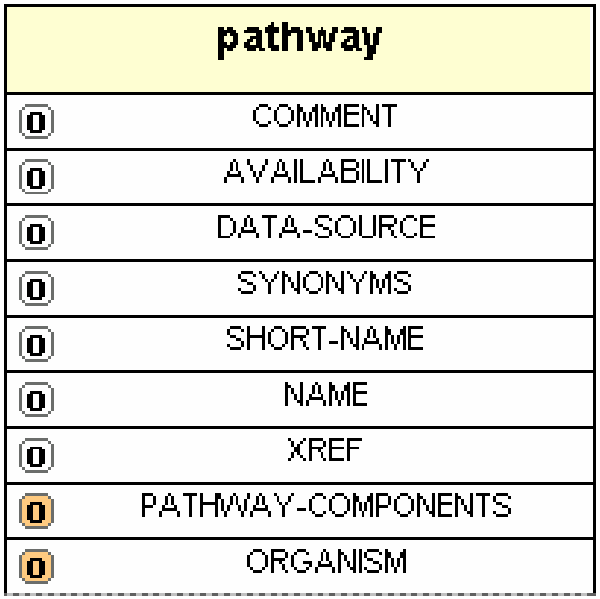
BioPAX Ontology: Root

- Root class: Entity
  - Any concept referred to as a discrete biological unit when describing pathways. This is the root class for all biological concepts in the ontology, which include pathways, interactions and physical entities



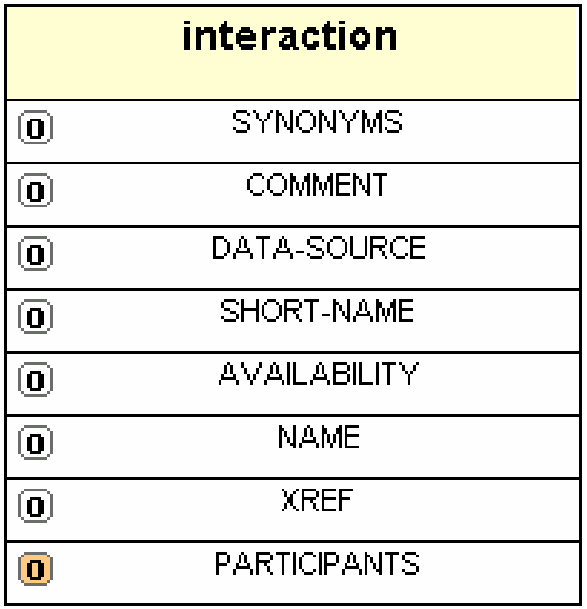
BioPAX Ontology: Pathway

- Pathway
  - An entity that consists of a set of interactions. A pathway is a series of molecular interactions and reactions, often forming a network, which biologists have found useful to group together for organizational, historic, biophysical or other reasons.



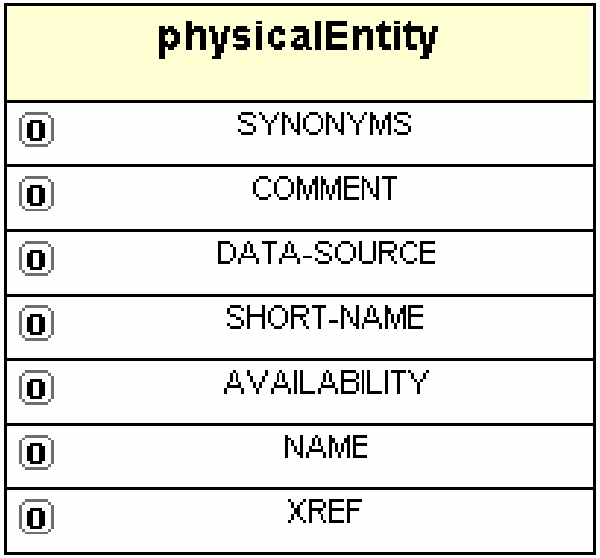
BioPAX Ontology: Interaction

- Interaction
  - An entity that defines a single biochemical interaction between **two or more** entities. An interaction cannot be defined without the entities it relates.

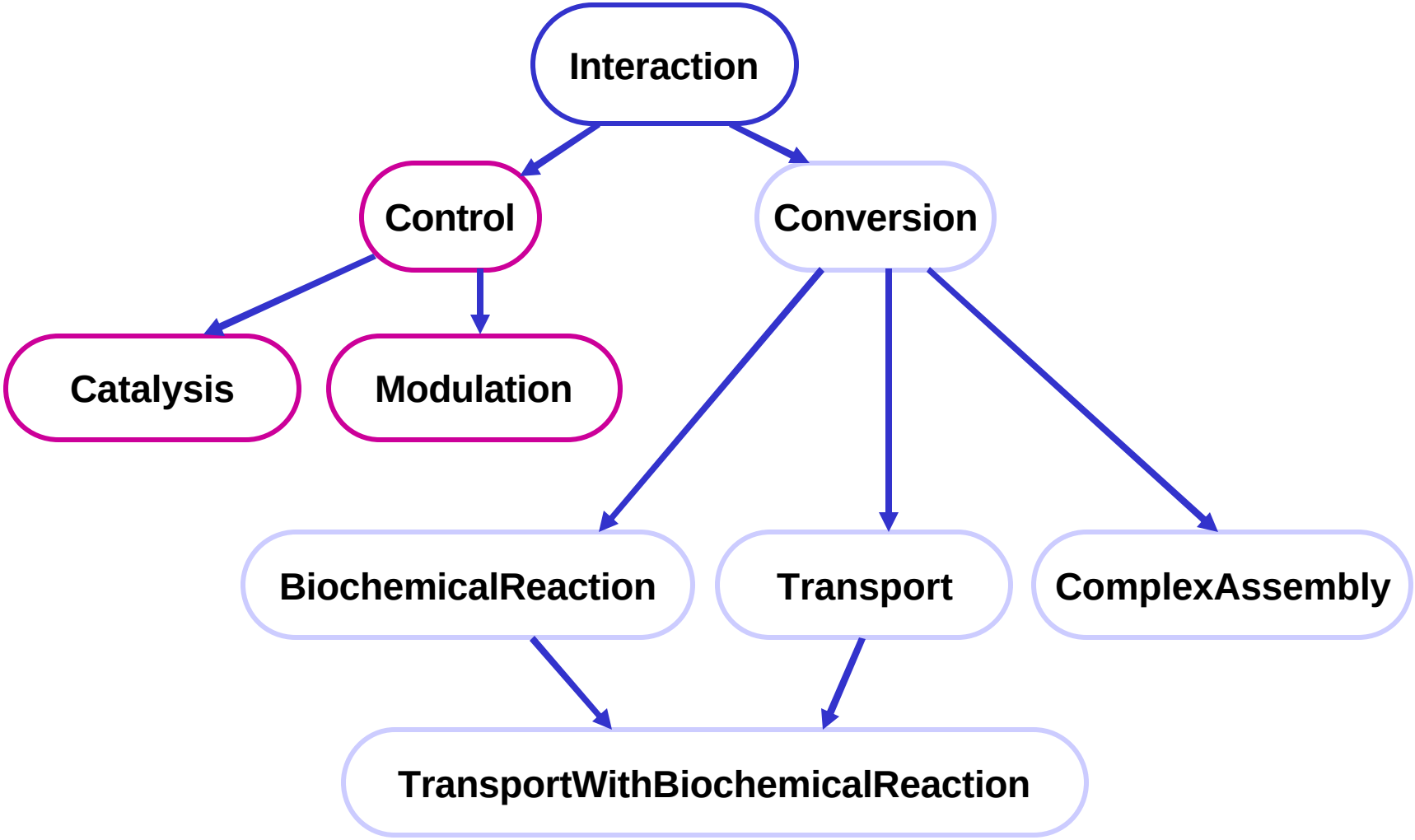


BioPAX Ontology: Parts

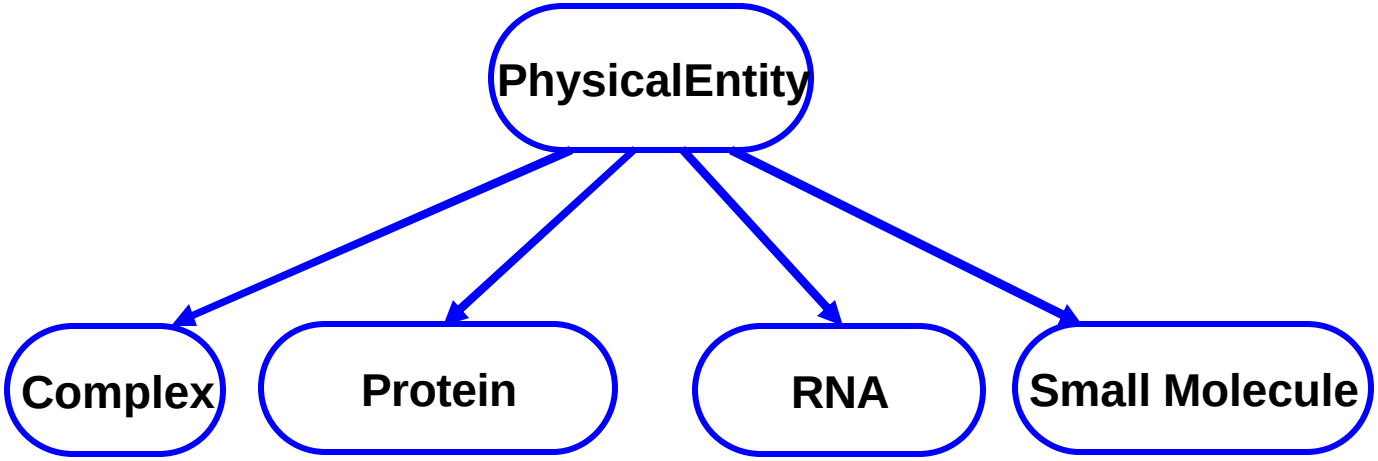
- Physical Entity
  - Any concept referred to as a discrete biological unit when describing pathways. This is the root class for all biological concepts in the ontology, which include pathways, interactions and physical entities



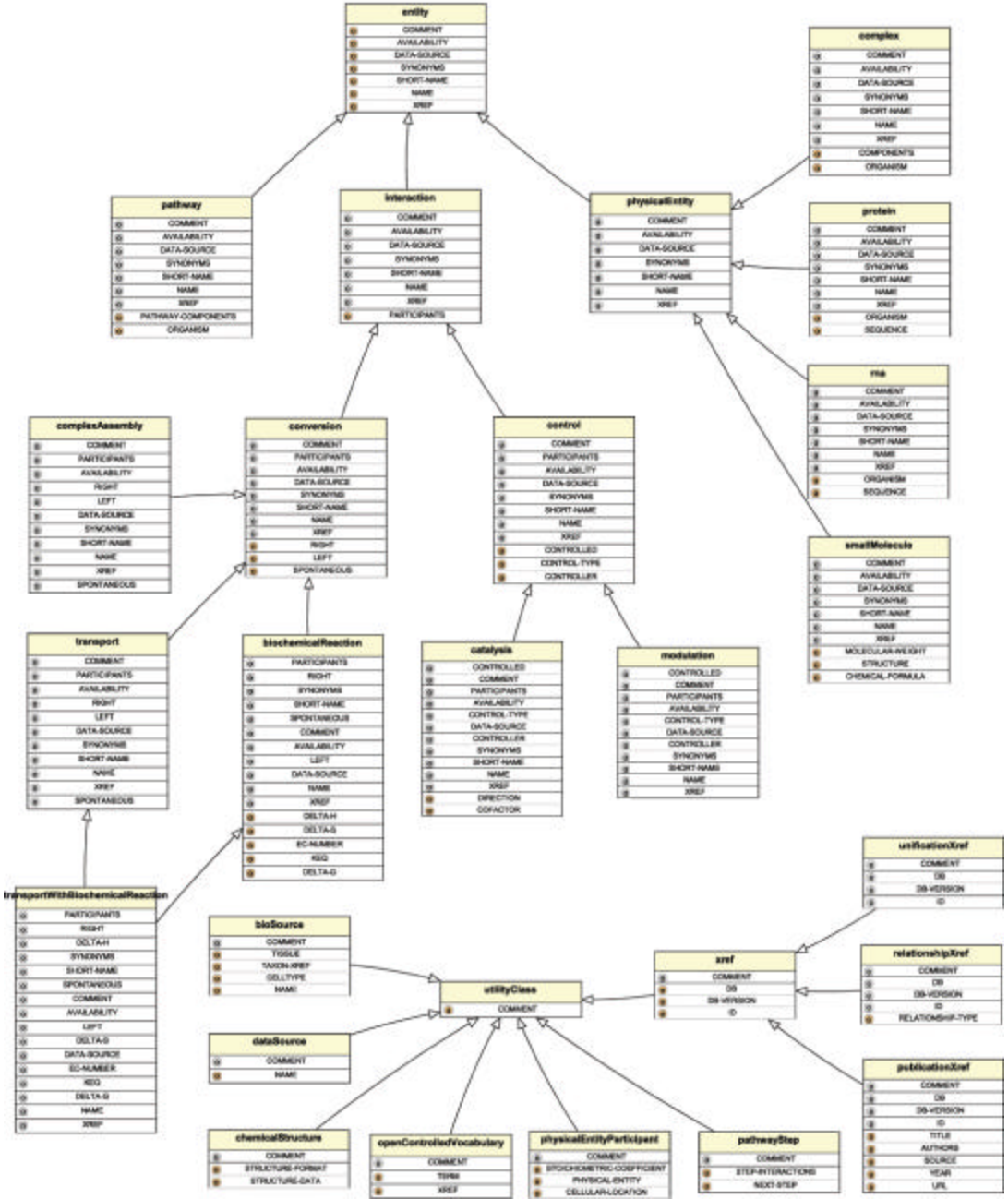
BioPAX Ontology: Interactions



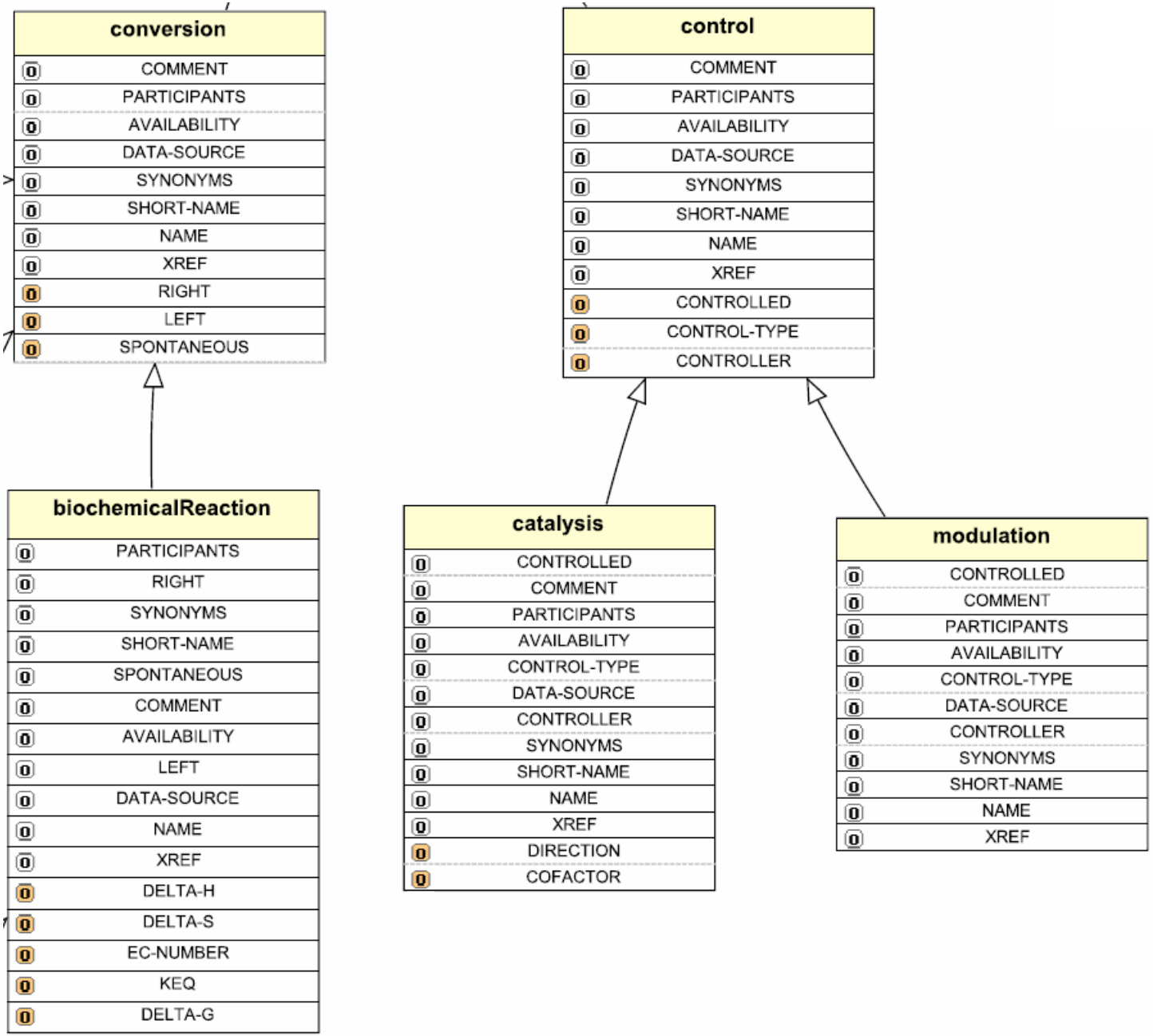
BioPAX Ontology: Physical Entities



BioPAX Ontology



BioPAX Ontology



Organizational Structure

- Small core group advancing standard
- Increased representation from mailing lists and subgroups
- Cost paid by participants/DOE workshop grant
- Special interests have subgroups
  - Core group member + outside experts
  - Tackle specific challenges

Current Status

- Initial meeting Nov. 2002
- Version 1.0 of Level 1 Ontology will be released July.7.2004
- Translating records from major DBs to BioPAX
- BOF meeting at ISMB 2004 (Glasgow) to discuss future directions
- Future directions: support regulatory and genetic pathways

BioPAX Supporting Groups

Groups

- Memorial Sloan-Kettering Cancer Center: G. Bader, M. Cary, C. Sander
- SRI Bioinformatics Research Group: P. Karp, S. Paley, J. Pick
- University of Colorado Health Sciences Center: I. Shah
- BioPathways Consortium: J. Luciano, E. Neumann, A. Regev, V. Schachter
- Argonne National Laboratory: N. Maltsev, E. Marland
- Samuel Lunenfeld Research Institute: C. Hogue
- Harvard Medical School: E. Brauner, D. Marks, A. Regev
- NIST: R. Goldberg
- Stanford: T. Klein
- Columbia: A. Rzhetsky
- Dana Farber Cancer Institute: J. Zucker

Collaborating Organizations:

- Proteomics Standards Initiative (PSI)
- Systems Biology Markup Language (SBML)
- CellML
- Chemical Markup Language (CML)

Databases

- BioCyc (www.biocyc.org)
- BIND (www.bind.ca)
- WIT (wit.mcs.anl.gov/WIT2)
- PharmGKB (www.pharmgkb.org)

Grants

- Department of Energy (Workshop)